

Rec'd PTO 08 FEB 2005

## 同步数字体系网络传输数据业务的流量控制方法

### 技术领域

- 5 本发明涉及数字信息的传输,尤其涉及一种用于通过同步数字体系 SDH 网络传输数据业务时的流量控制方法。

### 背景技术

- 随着同步数字体系 (SDH) 网络的快速发展, EoS (Ethernet over SDH/SONET: 通过 SDH/SONET 网络传输以太网业务) 技术也被更加广泛
- 10 泛的采用, 如何高效地通过 SDH 网络实现数据应用就显得更加重要, 在城域网的以太网传输应用中, 因为站点之间的距离可以达到一两百公里, 加上目前的 EoS 中的映射处理一般使用虚级联工作方式, 两站点之间的延时可能达到几毫秒甚至数十毫秒级, 再考虑到以太网业务的突发特性, 即瞬时流量很高, 一般情况下, 瞬时流量都可能超过预
- 15 先分配的带宽。

- 目前一种常用的流量控制机制是 CAR, (Committed Access Rate: 预定的接入速率), 它的实现原理是当网络发生拥塞时, 对于超出承诺速率的数据进行丢弃, 在丢弃时可能使用优先级概念。这种控制机制的缺点是一般都会造成大量的丢包现象, 在一个常用的 TCP/IP 运
- 20 用中, 大量丢包将严重降低网络性能, 特别是对于最广泛使用的 TCP/IP 协议, 丢包将造成其应用性能大幅度下降, 例如, 在 TCP 应用中, 使用 10M/b 的以太网设备, 如果超时重传定时器 AckTime 的值

等于 5 秒, 对于 1500Bytes 长的以太网帧, 1%的帧丢失率 FLR 将导致操作性能降低 98%。

有些网络传输中, 将用户数据设备产生的 PAUSE 帧传送到对端的用户数据设备, 但是传输过程中, 因为传输设备本身不会产生流控信息, 仅仅是将其它设备产生的 PAUSE 帧传递过去, 在短距离内可以使用, 不适合在长距离传输中使用; 另外, 如果有一端的设备不支持 PAUSE 帧, 这种方法不能使用。

如图 1 所示, EoS 设备由三部分组成: 以太网接入部分, 包括物理层接口 PHY 和媒体访问控制 MAC; 封装/解封装部分, 其封装方式可以是: LAPS (Link Access Protocol -SDH, 链路接入处理协议)、HDLCL (High level Data Link Control: 高级数据链路控制规程)、GFP (General Framing Procedure, 通用成帧处理) 等; 以及映射/解映射部分, 映射方式通常是虚级联、级联。

现假设 SDH 网络分配的带宽是 1 个 VC3, 即 45M, 当用户数据设备 A 的发送速率超过 1 个 VC3 时, 在现有的处理技术中, 因为映射 A 部分的流量恒定且为 1 个 VC3, 于是, 多出的以太网帧在封装 A 部分堆积, 当封装 A 部分的缓存耗尽后, 造成 MAC A 的数据不能继续发送到封装 A 部分, 于是, 当 PHY+MAC A 在允许以太网流控时, 向用户设备发 PAUSE 帧 (全双工) 或反压 (半双工), 即图 1 中的虚线 1 (如果禁止以太网流控, MAC A 丢弃多出的以太网帧, 以太网流控请参考 IEEE 802.3x 文档), 这将使设备 A 的发送速率减低到 1 个 VC3 以下而不会丢包, 从而以免网络性能过度降低。

当设备 A 发送速率低于 1 个 VC3，但因为某种原因使设备 B 的接收速率小于设备 A 的发送速率，例如设备 B 下挂多个 10M 设备，但是只有 1 个在工作，从设备 A 直到 EoS 设备 B 的发送端均不会阻塞，但是设备 B 本身将阻塞，于是，按照 802.3x 的实现方式，设备 B 将向和它直接相连的端口发送 802.3x 定义的流控信息，即图 1 中的虚线 2，可以是 PAUSE 帧或者反压信号，此流控信息按照 802.3x 的实现方式，将被 MAC B 终结，同时 MAC B 将停止向设备 B 发送数据，不会传递到设备 A，设备 A 继续按照原来的流量发送数据，于是，造成来自设备 A 的以太网帧在解封装 B 或者 MAC B 处堆积，当这两处的缓存耗尽后，EoS#B 必定丢包，这样，虽然设备 A 和设备 B 都是支持 802.3x 以太网流控的设备，但是通过传输设备 EoS 后，实际上不能避免丢包现象，这就说明，按照 802.3x 实现的现有技术方案不能通过 EoS 设备真正起到以太网的流控作用。

### 发明内容

15 本发明的目的在于提供一种高性能的用于通过同步数字体系 SDH 网络传输数据业务时的流量控制方法，实现以太网数据通过 SDH 网络的无丢失透明传输。

本发明所采用的方法为：一种同步数字体系网络传输数据业务的流量控制方法，包括以下步骤：

20 A. 在同步数字体系 (SDH) 网络的数据传输中，EoS (Ethernet over SDH/SONET: 以太网在 SDH/SONET 中传输) 设备的封装部分，根据帧缓存的使用情况，产生并封装流控协议帧 (LFP: Line Flow-control

Protocol, 线路流量控制协议);

B. 将流控协议 (LFP) 帧作为普通数据帧映射进 SDH 净荷中, 传输到对端设备;

C. 对端设备将 SDH 净荷解映射处理后, 由 EoS 处理设备的解封部分进行流控协议 (LFP) 帧识别, 并解释执行流控协议 (LFP) 帧中携带的流量控制信息。

所述产生并封装流控协议 (LFP) 帧的步骤进一步包括:

A1. 在所述的 EoS 设备的上行方向, EoS 处理设备的封装部分随时监测缓存内的数据量, 产生相应控制域的流控协议 (LFP) 帧, 当高于高水线时, 定时重复发送控制域为停止发送的流控协议 (LFP) 帧; 当低于低水线时, 定时重复发送控制域为启动发送的流控协议 (LFP) 帧; 若处于高、低水线之间时, 停止发送流控协议 (LFP) 帧;

A2. 该 LFP 帧被插入到封装数据队列的最前面, 优先作封装处理: 如果当前没有封装以太网帧, 立即封装该 LFP 帧, 否则, 等待当前的以太网帧封装完成后, 立即封装该 LFP 帧。

所述的流控协议 (LFP) 帧的载体可使用标准的 802.3x 定义的 PAUSE 帧结构。

本发明的有益效果为: 在本发明中, 通过使用流控协议 LFP 帧, 在 EoS 设备之间通过 SDH 网络相互传递流量控制信息, 传输设备自身产生流控协议 LFP 帧, 进行网络流量控制信息的相互传递, 以实现突发方式的以太网数据在周期性传输的 SDH 网络中的无丢包传送, 实现以太网数据通过 SDH 网络的无丢失传输, 通过无丢失传输以太网数据

来大幅度提高数据应用的性能。本发明能够在 EoS 设备之间传递流控信息，而且能够适用于 EoS 下挂全双工或者半双工设备，例如可以一端是全双工设备，另一端是半双工设备，或者两端都是半双工设备；EoS 处理设备随时监测缓存内的数据量，可以防止流制协议 LFP 帧因为传输出错等意外原因而引起的流量控制失败，因此，本发明是一种高性能的网络信息流控方法，适用性强，工作可靠。

### 附图说明

图 1 为现有技术信息流控示意图；

图 2 为本发明流控协议 LFP 处理示意图。

### 10 具体实施方式

下面根据附图和实施例对本发明作进一步详细说明，下面的框图描述中，一般情况下，以从 A 到 B 方向传输为例，从 B 到 A 的处理方法相同。

在本发明中，在通过同步数字体系 SDH 网络的数据传输中，EoS 设备之间通过流控协议 LFP 帧相互传递流量控制信息，流控协议 LFP 帧的载体可使用标准的 802.3x 定义的 PAUSE 帧结构。

EoS 的映射部分在做 SDH 映射时将以太网帧和流控协议 LFP 帧都做为普通数据帧处理，流控协议 LFP 帧的控制域为 0x0FFFFH 或 0x0H，采用 Xon/Xoff 的开/关控制方式。

另外，因为流控协议 LFP 帧在传输过程中有延时，需要使用适量的数据缓存器件以弥补传输延时；流控协议 LFP 帧的实现是基于带内管理的方式，即 LFP 帧和普通以太网数据帧采用同样的传输路径。

在本发明中,如图 2 所示,在上行方向,即系统方向,如图 2 中的 EoS#A 到 SDH 网络方向,以 EoS#A 为例,EoS 处理设备的封装部分,根据帧缓存的高低水准值,(该水准反映出缓存的使用情况),产生控制域为 0x0FFFFH 或 0x0H 的流控协议 LFP 帧,该流控协议 LFP 帧在 EoS 处理的映射部分中被看成普通以太网数据帧,但是为了提高 LFP 的性能,要求优先将流控协议 LFP 帧发送到 EoS 处理的映射部分;在下行方向,即到用户设备方向,如图 2 中的 SDH 到设备 EoS#B 方向,以 EoS#B 为例,EoS 处理设备的解封装部分进行流控协议 LFP 帧识别,当发现是流控协议 LFP 帧时,解释执行该流控协议 LFP,从而实现流控协议 LFP 协议。

当 EoS#A 下挂的用户设备 A 的接收端,即图 2 中的 AR,发生阻塞时,因为解映射 A 照常工作,缓存 A 中的数据量将增加,当达到设定的高位水准时,解封装 A 产生控制域为 0x0FFFFH 的流控协议 LFP 帧,该流控协议 LFP 帧被插入到封装 A 数据队列的最前面,优先作封装处理,其具体过程是如果封装 A 当前没有封装以太网帧,立即封装该流控协议 LFP 帧,否则,等待当前的以太网帧封装完成后,立即封装该流控协议 LFP 帧,再经过映射处理后,经过 SDH 传输网络、解映射 B 到达解封装 B,解封装 B 在识别出是流控协议 LFP 帧后,有两种处理方式:

1、LFP 透传方式:在网络延时很小,两站点间的距离很近的情况下采用此种方式,解封装 B 将根据本端下挂的数据设备是否是全双工模式来解释执行该流控协议 LFP 帧:

如果 EoS 设备下挂的用户数据设备工作于全双工模式，不需要解释流控协议 LFP 帧，直接将流控协议 LFP 帧发送到用户数据设备 B，其具体过程是如果当前解封装 B 没有向以太网接入部分 BR 发送数据，则立即发送此流控协议 LFP 帧到以太网接入部分 BR，否则，等当前以太网数据帧发送完成后，立即发送此流控协议 LFP 帧，此时流控协议 LFP 帧不经过缓存 B 以提高 LFP 协议的性能；

如果 EoS 设备下挂用户数据设备工作于半双工模式，需要解释流控协议 LFP 帧中控制域的值，具体实现是当流控协议 LFP 帧中控制域控制域 = 0x0H 时，撤销反压控制信号，否则，发送反压控制信号，让 EoS#B 下挂的用户设备 B 检测到冲突，从而停止发送数据到封装 B 中。

2、LFP 再生方式：解封装 B 将解释执行该流控协议 LFP 帧，当流控协议 LFP 帧中控制域的值不是 0x0H，即 LFP 禁止发送，于是封装 B 停止工作，即停止接收以太网接入部分 BT 送过来的数据，如果是解封装 B 自身产生的流控协议 LFP 帧，还是需要通过封装 B 送到映射 B 中。这样，来自以太网接入部分 BT 的数据将在“封装 B”处堆积，并通过 802.3x 定义的以太网流控协议，最终将使用户设备 B 停止发送以太网帧，即 BT 点没有数据进入；当流控协议 LFP 帧中控制域的值 = 0x0H，即允许发送，封装 B 按照正常的方式工作，将以太网接入部分 BT 进入的以太网帧经过封装送到映射 B 中。

该方式下，不要求 SDH 网络两端 EoS 的外挂设备均工作于全双工模式，只要用户设备支持 802.3x 定义的标准流控操作即可，可以一

端工作于全双工模式，另一端是半双工模式；或者两端都是半双工模式。

当 EoS#B 下挂的用户设备 B 停止向 EoS#B 发送数据后，即 BT 点没有以太网帧输入，EoS#A 的缓存 A 中数据将逐渐减少，当达到低位  
5 水线时，解封装 A 产生控制域为 0x0H 的流控协议 LFP 帧，该流控协议 LFP 帧按上述步骤优先送到 EoS#B 的解封装 B 中解释执行，结果使 EoS#B 下挂的用户设备 B 重新发送数据。

为了防止流控协议 LFP 帧因为传输出错等意外原因而引起的 LFP 控制失败，EoS 处理设备随时监测缓存内的数据量，当高于高水线或  
10 低于低水线时，定时重复发送相应控制域的流控协议 LFP 帧；若处于高、低水线之间时，停止发送流控协议 LFP 帧，当高于高水线时，应该定时重复发送控制域为 0x0FFFFH 的流控协议 LFP 帧；当低于低水线时，应该定时重复发送控制域为 0x0H 的流控协议 LFP 帧。

高、低水线和定时重发时间可以设置，这 2 个定时重发时间可以  
15 不同，建议发送控制域为 0x0FFFFH 的流控协议 LFP 帧的重发时间间隔不要过长；而发送控制域为 0x0H 的流控协议 LFP 帧的重发时间间隔不要过短，以避免流控协议 LFP 帧过度占用有效带宽。

目前 LFP 帧使用的格式和 802.3x 定义的标准 PAUSE 帧相同，这是为了简化 LFP 帧透传方式的设计，不是必须的，也可以使用其它格  
20 式，但在 LFP 帧透传方式的接收端需进行格式转换；LFP 帧再生方式因为 LFP 帧只是在 EoS 设备之间传递，所以可以不用转换格式。

本发明同样适用于 SONET 网络，其原理和控制过程基本一致，此



处不再赘述。

## 权 利 要 求

1. 一种同步数字体系网络传输数据业务的流量控制方法, 其特征在于, 包括以下步骤:

- A. 在同步数字体系(SDH)网络的数据传输中, EoS(Ethernet over  
5 SDH/SONET: 以太网在 SDH/SONET 中传输)设备的封装部分根据帧缓存的使用情况, 产生并封装流控协议(LFP)帧;
- B. 将流控协议(LFP)帧作为普通数据帧映射进 SDH 净荷中, 传输到对端设备;
- C. 对端设备将 SDH 净荷经解映射处理后, 由 EoS 处理设备的解封  
10 装部分进行流控协议(LFP)帧识别, 并解释执行流控协议(LFP)帧中携带的流量控制信息。

2. 根据权利要求 1 所述的同步数字体系网络传输数据业务的流量控制方法, 其特征在于: 产生并封装流控协议(LFP)帧的步骤包括:

- 15 A1. 在所述的 EoS 设备的上行方向, EoS 处理设备的封装部分随时监测缓存内的数据量, 产生相应控制域的流控协议(LFP)帧, 当高于高水线时, 定时重复发送控制域为停止发送的流控协议(LFP)帧; 当低于低水线时, 定时重复发送控制域为启动发送的流控协议(LFP)帧; 若处于高、低水线之间时, 停止发送流控协议(LFP)帧;
- 20 A2. 所述 LFP 帧被插入到封装数据队列的最前面, 优先作封装处理: 如果当前没有封装以太网帧, 立即封装该 LFP 帧, 否则, 等待当前的以太网帧封装完成后, 立即封装该 LFP 帧。

3. 根据权利要求 1 或 2 所述的同步数字体系网络传输数据业务的流量控制方法，其特征在于：所述的流控协议（LFP）帧的载体使用标准的 802.3x 定义的 PAUSE 帧结构。

4. 根据权利要求 2 所述的同步数字体系网络传输数据业务的流量控制方法，其特征在于：步骤 A1 中缓存内的数据量高于高水线时，定时重复发送控制域为 0x0FFFFH 的流控协议（LFP）帧；当低于低水线时，定时重复发送控制域为 0x0H 的流控协议（LFP）帧，控制域采用 Xon/Xoff 的开/关控制方式。

5. 根据权利要求 1 所述的同步数字体系网络传输数据业务的流量控制方法，其特征在于：所述的解封装在识别出是流控协议（LFP）帧后，对流控协议（LFP）帧作出处理时，可采用流控协议（LFP）帧的再生方式：

当 LFP 帧中的控制域不是启动发送时，即 LFP 禁止发送，于是封装部分停止工作，来自以太网接入部分的数据将在封装部分处堆积，并通过以太网的流控协议（LFP）帧，最终使用户数据设备停止发送以太网帧；

当流控协议（LFP）帧中控制域为启动发送时，即允许发送，封装按照正常的方式工作，将以太网接入部分进入的以太网帧经过封装送到映射中。

6. 根据权利要求 1 所述的同步数字体系网络传输数据业务的流量控制方法，其特征在于：所述的解封装在识别出是流控协议（LFP）帧后，对流控协议（LFP）帧作出处理时，可采用流控协议（LFP）透

传方式,解封装将根据本端下挂的用户数据设备是否支持全双工模式来解释执行该流控协议(LFP)帧:

如果 EoS 设备下挂的用户数据设备工作于全双工模式,不需要解释流控协议(LFP)帧,直接将流控协议(LFP)帧发送到用户数据设备;

如果 EoS 设备下挂的用户数据设备工作于半双工模式,需要解释 LFP 帧中控制域的值,当流控协议(LFP)帧中控制域为启动发送时,撤销反压控制信号,否则,发送反压控制信号,让 EoS 设备下挂的用户数据设备检测到冲突,从而停止发送数据。

- 10        7. 根据权利要求 5 或 6 所述的同步数字体系网络传输数据业务的流量控制方法,其特征在于:当 EoS 设备下挂的用户数据设备停止向 EoS 设备发送数据后,对端 EoS 设备的缓存中数据将逐渐减少,当达到低位水线时,对端解封装部分产生控制域为启动发送的流控协议(LFP)帧,该 LFP 帧优先送到本端的解封装部分解释执行,结果使
- 15    本端 EoS 下挂的用户数据设备重新发送数据。

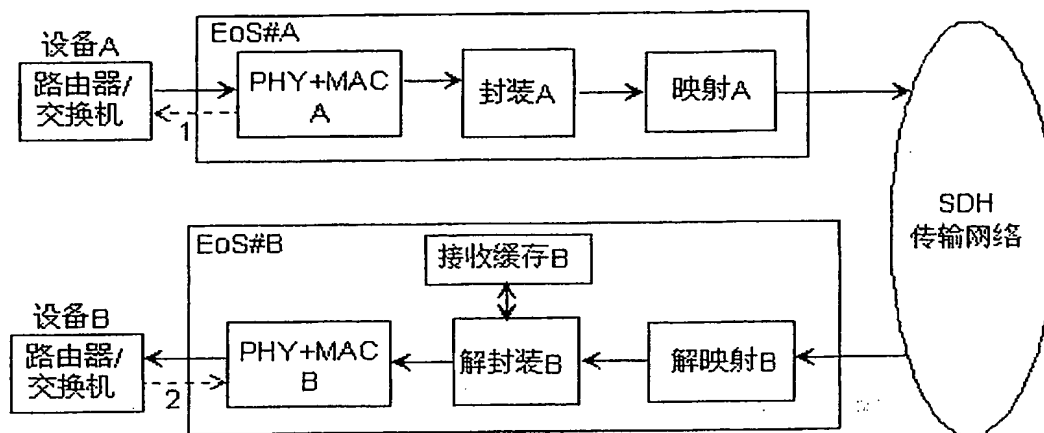


图 1

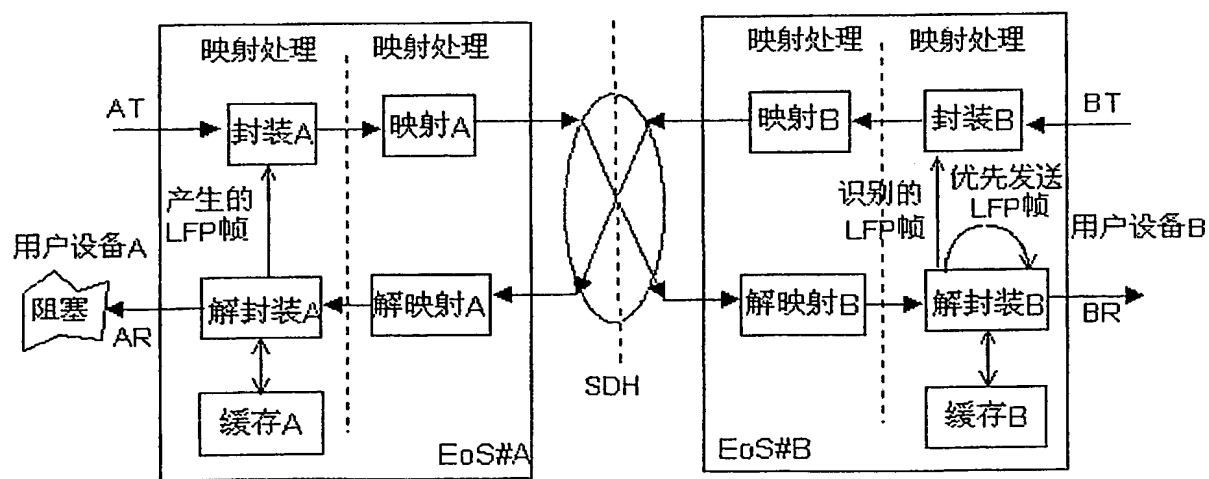


图 2

**This Page Blank (uspto)**

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/ CN03/00643

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04L 12/28

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H04L12/24, H04Q11/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, PAJ: ENCAPSULATE, FLOW, CONTROL, PROTOCOL, FRAME, SDH, CONGESTION

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN, A, 1353367, (SUANTONG DIGITAL TECHNOLOGY RES CENT CO), 12 JUNE 2002, page 2, line 3—page 7, line 14, fig.1—5	1—7
A	CN, A, 1333965, (TELEFONAKTIEBOLAGET ERICSSON L M), 30 Jan. 2002, page 3, line 29—page 14, line 3	1—7
A	WO, A1, 0131970, (E-CELL TECHNOLOGIES), 3 May 2001, page 1, paragraph 1—page 18, paragraph 3	1—7
A	EP, A2, 1006751, (NORTEL NETWORKS CORP), 7 June 2000, page 1, line 43—page 11, line 33	1—7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

\* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

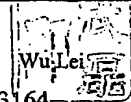
“&” document member of the same patent family

Date of the actual completion of the international search  
18 Sep. 2003

Date of mailing of the international search report  
30 OCT 2003 (30.10.03)

Name and mailing address of the ISA/CN  
6 Xitucheng Rd., Jimen Bridge, Haidian District,  
100088 Beijing, China  
Facsimile No. 86-10-62019451

Authorized officer



Telephone No. 86-10-62093164

**This Page Blank (uspto)**



# 国际检索报告

国际申请号  
PCT/ CN03/00643

## A. 主题的分类

IPC7: H04L 12/28

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

## B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

IPC7: H04Q11/04, H04L12/44

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

CNPAT: 封装, 流量, 控制, 协议, 帧, SDH

WPI, EPODOC, PAJ: ENCAPSULATE, FLOW, CONTROL, PROTOCOL, FRAME, SDH, CONGESTION

## C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求编号
A	CN, A, 1353367, (北京算通数字技术研究中心有限公司), 2002年6月12日, 说明书第2页第3行至第7页第14行, 图1-5	1-7
A	CN, A, 1333965, (艾利森电话股份有限公司), 2002年1月30日, 说明书第3页第29行至第14页第3行	1-7
A	WO, A1, 0131970, (E-CELL TECHNOLOGIES), 2001年5月3日, 说明书第1页第1段至第18页第3段	1-7
A	EP, A2, 1006751, (NORTEL NETWORKS CORP), 2000年6月7日, 说明书第1页第43行至第11页第33行	1-7

☐ 其余文件在 C 栏的续页中列出。

☐ 见同族专利附件。

### \* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件

“E” 在国际申请日的当天或之后公布的在先的申请或专利

“L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理

“X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性

“&” 同族专利成员的文件

国际检索实际完成的日期

18. 9 月 2003 (18. 09. 03)

国际检索报告邮寄日期

3 0. 10 月 2003 (3 0. 1 0. 03)

国际检索单位名称和邮寄地址

ISA/CN

中国北京市海淀区西土城路 6 号(100088)

传真号: 86-10-62019451

授权官员

电话号码: 86-10-62093164



**This Page Blank (uspto)**